

Multiplicative Noise and Non-Gaussianity: A Paradigm for Atmospheric Regimes?

PHILLIP SURA, MATTHEW NEWMAN, CÉCILE PENLAND, AND PRASHANT D. SARDESHMUKH

NOAA-CIRES Climate Diagnostics Center, University of Colorado, Boulder, Colorado

(submitted to *J. Atmos. Sci.* January 2004)

ABSTRACT

Atmospheric circulation statistics are not strictly Gaussian. Small bumps and other deviations from Gaussian probability distributions are often interpreted as implying the existence of distinct and persistent nonlinear circulation regimes, with potentially higher than average levels of predictability. In this paper it is shown that such deviations from Gaussianity can, however, also result from linear stochastically perturbed dynamics with multiplicative (i.e., state-dependent) noise statistics, but are generally associated with much lower levels of predictability. Multiplicative noise is often identified with state-dependent variations of stochastic feedbacks from unresolved system components, and may be treated as stochastic perturbations of system parameters. It is shown that including such perturbations in the damping of large-scale linear Rossby waves can lead to deviations from Gaussianity very similar to those in the joint probability distribution of the first two Principal Components (PCs) of observed weekly-averaged 750 hPa streamfunction data for the past 52 winters. A closer examination of the Fokker-Planck probability budget in the plane spanned by these two PCs shows that the observed deviations from Gaussianity can be modeled with multiplicative noise, and are unlikely the results of slow nonlinear interactions of the two PCs. It is concluded that the observed non-Gaussian probability distributions do not necessarily imply the existence of persistent nonlinear circulation regimes, and are entirely consistent with those expected for a linear system perturbed by multiplicative noise.

1. Introduction

Predictability of the atmosphere is limited by the chaotic interaction of weather. There are many different ways of defining predictability; for example, we may consider the doubling time of initial uncertainty. It is of considerable forecast interest to find those atmospheric states for which this error doubling time is particularly long.

Attention has thus been paid to the possible existence of persistent or quasi-stationary atmospheric flow regimes, defined as recurrent large-scale atmospheric flow configurations that persist longer than individual weather systems (e.g., Pandolfo 1993). It is hoped that these states may be inherently more predictable and/or that they may provide initial atmospheric states from which more reliable forecasts are made. The idea of weather regimes, in terms of midlatitude cyclone tracks over Europe, was first proposed by van Bebber (1891). The more modern phenomenological notion of midlatitude atmospheric flow regimes, or *Großwetterlagen*, was introduced by Baur et al. (1944) and Baur (1947) for use in making statistical long-range weather forecasts. Dynamical theories of multiple equilibria (blocked and

zonal regimes) due to nonlinear wave-mean flow interactions, first proposed by Charney and DeVore (1979), Wiin-Nielsen (1979), and Hart (1979), spurred a renewed effort to explore the existence of regimes in observed midlatitude flows (e.g., Sutera 1986; Hansen and Sutera 1986; Mo and Ghil 1988; Molteni et al. 1990; Kimoto and Ghil 1993a,b; Cheng and Wallace 1993; Corti et al. 1999; Smyth et al. 1999; Monahan et al. 2001, and many others). It should be noted that these observational studies generally did not find regime behavior as pronounced as in the theoretical and simple model studies.

The term “regime” can be nonspecific in common usage. Some stable linear systems can have “regimes” if what is meant is rapid singular-vector growth followed by a quasi-stationary state (e.g., Farrell and Ioannou 1996). Such observed atmospheric behavior can be well simulated by relatively low order multivariate linear models (Cash and Lee 2001; Winkler et al. 2001). In fact, to the extent that these models are good forecast models of weekly averages (Newman et al. 2003), they suggest an important constraint on the nature of nonlinearity in the atmosphere; we will return to this point later.

A stable linear system driven by Gaussian forcing will result in Gaussian statistics. Therefore, studies exploring potential regime behavior resulting from nonlinear dynamics have focused upon the non-Gaussianity of the Probability Distribution Function (PDF) representing all possible atmospheric states. Ideally, one would like to

Corresponding author address: Philip Sura, NOAA-CIRES Climate Diagnostics Center, Mail Code R/CDC, 325 Broadway, Boulder CO 80305-3328.
E-Mail: Phillip.Sura@noaa.gov

find significant multiple peaks in these multivariate PDFs. However, due to the limited data record, observational studies of climate regimes more typically examine the bivariate PDF describing the two leading Empirical Orthogonal Functions (EOFs) of an appropriate atmospheric variable (e.g., Mo and Ghil 1988; Molteni et al. 1990; Kimoto and Ghil 1993a,b; Corti et al. 1999; Smyth et al. 1999; Weisheimer et al. 2001). Moreover, the PDFs of observed large scale atmospheric circulation anomalies do not show any statistically significant multimodality. Rather, PDF “inhomogeneities” (statistically significant deviations from bivariate Gaussianity) are often interpreted as multiple Gaussian regimes (Smyth et al. 1999).

The most common explanation for extratropical climate regimes is that they are induced by nonlinearities in the slow manifold of the equations governing atmospheric dynamics (e.g., Legras and Ghil 1985; Yoden 1985a,b; Ghil and Childress 1987; DeSwart 1988; Itoh and Kimoto 1996, 1997, 1999; Ghil and Robertson 2002, and many others). It is this slow process that gives hope for potential long-range predictability. Of course, slowly evolving external non-Gaussian forcing, such as that produced by anomalous tropical convection due to the Madden-Julian oscillation (MJO) and El Niño-Southern Oscillation (ENSO), could also produce a non-Gaussian response in the extratropics even if extratropical dynamics are entirely linear.

Regimes that result from either or both of these mechanisms might indeed be more predictable than other atmospheric states. There is, however, a third possibility: regimes could result not from slow processes but rather from the fast (that is, rapidly decorrelating) nonlinearities of the dynamical system. This arises because, given the very high number of degrees of freedom in the atmosphere and the consequent plethora of nonlinear subsystems, it is highly likely that there exists chaos sufficiently disordered as to make the application of the Central Limit Theorem (e.g., Khasminskii 1966; Papanicolaou and Kohler 1974; Majda et al. 1999, 2003) valid at medium-range timescales (about a week to months). In this case, the fast nonlinearities may be approximated as state-dependent, or multiplicative, stochastic noise that is inherently unpredictable. [For state-independent, or additive, noise this was first noted in a climate context by Hasselmann (1976).]

That apparently similar regimes can be induced by either slow or fast nonlinearities is illustrated in Fig. 1 (a more quantitative discussion is given in section 3). Consider a double potential well as a simple model of two “regimes” resulting from nonlinearities of the slow manifold. Systems trajectories kicked from one potential well to the other by additive noise will result in a bimodal PDF. This is not, however, the only dynamical system which can produce such a PDF. Consider instead a linear system, represented by a monomodal deterministic potential, in which state trajectories are kicked around by multiplicative noise. If the noise is relatively strong in the center of the monomodal potential, and rel-

atively weak at the edges, then this system will also have a bimodal PDF. Thus the same bimodal PDF can result from either a slow (deterministic) nonlinear dynamical system or a fast (stochastic) nonlinear dynamical system. As we will see, however, the predictability of these two systems is very different.

In this paper we will use both simple models and observational analysis to develop the hypothesis that atmospheric non-Gaussian regimes may be due to multiplicative noise. First, in section 2, some principal results of stochastic dynamics are briefly reviewed. Then, in section 3, we discuss some simple examples that demonstrate how a given PDF can imply very different limits of predictability depending upon whether the PDF results from slow deterministic nonlinear evolution or fast stochastic evolution. In section 4, stochastic perturbations of the linear damping and the zonal mean flow are included in a linear model of Rossby wave evolution, to assess the impact of multiplicative noise in a simple but meteorologically relevant setting. By explicitly solving the Fokker-Planck equation and the stochastic differential equation, we show that multiplicative noise in the frictional damping leads to intermittency and consequently a highly non-Gaussian distribution. These results provide one possible explanation for the non-Gaussian PDF found in the observational analysis presented in section 5, where it is shown the regime behavior in the leading EOFs of 750 hPa Northern Hemisphere streamfunction is consistent with that resulting from multiplicative noise. Finally, section 6 provides a summary and a discussion.

2. Stochastic Dynamics in a Nutshell

This section introduces a few basic ideas of stochastic dynamics used in this paper. More comprehensive treatises can be found in many textbooks (e.g., Gardiner 1985; Horsthemke and L  f  ver 1984, Paul and Baschnaegle 1999).

Consider the dynamics of an n -dimensional system whose state vector \mathbf{x} is governed by the stochastic differential equation (SDE)

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\eta \quad (1)$$

where the vector $\mathbf{A}(\mathbf{x})$ represents all slow processes and $\mathbf{B}(\mathbf{x})\eta$, with the matrix $\mathbf{B}(\mathbf{x})$ and the noise vector η , represents the stochastic approximation to the fast nonlinear processes. The stochastic components η_i are assumed to be independent Gaussian white noise processes:

$$\langle \eta_i(t) \rangle = 0, \langle \eta_i(t)\eta_j(t') \rangle = \delta_{ij}\delta(t-t') \quad (2)$$

where $\langle \dots \rangle$ denotes the averaging operator. The corresponding Fokker-Planck equation,

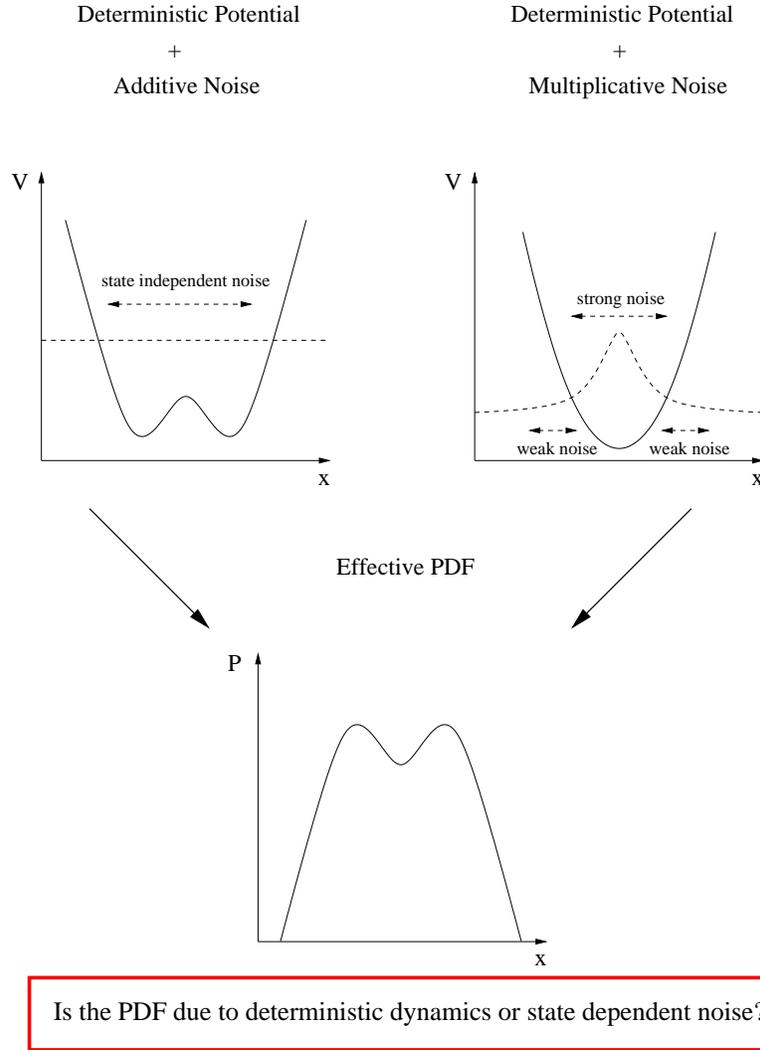


FIG. 1. A schematic sketch to illustrate the fundamental dynamical difference between deterministic and stochastically induced regimes. The effective PDF of a trajectory in a deterministic double-well potential driven by additive noise will be bimodal. The same effective PDF can be produced by a trajectory in a monomodal deterministic potential kicked around by multiplicative noise. Because of the larger noise amplitudes near the center of the monomodal potential, as compared to the strength of the noise right and left from it, the system's trajectory is more often found on either side of the central noise maximum. Thus, the PDF becomes bimodal. See appendix A for a mathematical formulation of this behavior.

$$\frac{\partial}{\partial t} p(\mathbf{x}, t) = -\sum_i \frac{\partial}{\partial x_i} \left[A_i + \alpha \sum_{j,k} \left(\frac{\partial B_{ik}}{\partial x_j} \right) B_{jk} \right] p(\mathbf{x}, t) \quad (3)$$

$$+ \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\mathbf{B}\mathbf{B}^T)_{ij} p(\mathbf{x}, t)$$

describes the conservation of the probability density $p(\mathbf{x}, t)$ of the system described by the SDE (1). Two different values of α yield two physically important stochastic calculi: the Itô ($\alpha=0$) and the Stratonovich calculus ($\alpha=1/2$). On the right hand side, the first term within square brackets describes the dynamics of the deterministic system and is called the deterministic drift. The second term within square brackets, which does not occur in Itô systems ($\alpha=0$), is called the noise-induced drift. The remaining term is associated with the diffusion

of the probability density by noise.

For a detailed discussion of stochastic integration and the differences between Itô and Stratonovich SDEs see for example Horsthemke and Léfèvre (1984), Gardiner (1985), or Penland (1996). The key point here is that the Stratonovich calculus is relevant for continuous physical systems, such as the atmosphere, in which rapidly fluctuating quantities with small but finite correlation times are approximated as white noise. Thus, simplified stochastic models constructed from atmospheric dynamical equations may assume Stratonovich calculus. However, if instead a stochastic model is indirectly estimated from observed discrete data, then the inferred drift will be the sum of the deterministic and the noise-induced drifts. In this case using the Itô framework may be preferable, where now $\mathbf{A}(\mathbf{x})$ represents not just the deterministic

drift but rather this sum, or the “effective drift”.

In principle, the deterministic and stochastic parts of (3) can be determined from data by using their statistical definitions (Siegert et al. 1998; Friedrich et al. 2000; Gradisek et al. 2000; Sura and Barsugli 2002; Sura 2003; Sura and Gille 2003):

$$\mathbf{A}(\mathbf{x}) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle \mathbf{X}(t + \Delta t) - \mathbf{x} \rangle \Big|_{\mathbf{x}(t) = \mathbf{x}} \quad (4)$$

$$\mathbf{B}(\mathbf{x})\mathbf{B}^T(\mathbf{x}) \quad (5)$$

$$= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle (\mathbf{X}(t + \Delta t) - \mathbf{x})(\mathbf{X}(t + \Delta t) - \mathbf{x})^T \rangle \Big|_{\mathbf{x}(t) = \mathbf{x}}$$

where $\mathbf{X}(t + \Delta t)$ is a solution (a single stochastic realization) of the SDE (1) with the initial condition $\mathbf{X}(t) = \mathbf{x}$ at time t . The data define a state space representing every observed value of \mathbf{x} . The effective drift and stochastic diffusion are estimated by replacing the theoretical limit $\Delta t \rightarrow 0$ with a finite-difference approximation. In practice, estimating $\mathbf{B}(\mathbf{x})\mathbf{B}^T(\mathbf{x})$ from discretely sampled data is prone to error, because Taylor expansions of stochastic terms are proportional to $\sqrt{\Delta t}$ and not proportional to Δt as are the deterministic terms (e.g., Sura and Barsugli 2002; Sura 2003).

When \mathbf{A} and \mathbf{B} are known, analytical solutions of the Fokker-Planck equation (3) for $p(\mathbf{x}, t)$ can only be found in limited cases (appendix B presents one such case). For more general cases, numerical methods must be used. In this paper (section 4) the semi-implicit Chang-Cooper method is implemented to solve (3) for $p(\mathbf{x}, t)$ (Chang and Cooper 1970; Park and Petrosian 1996). It employs a flux-conservative second-order accurate finite difference scheme, extended to multidimensional problems using the operator splitting method (Press et al. 1992; Park and Petrosian 1996). In section 4 a regular grid with a mesh size 0.1 and 200x200 grid points is used. The domain of computation is [-10:10,-10:10]. The Fokker-Planck equation is integrated until a steady state is reached. The initial distribution of the probability density function p is chosen to be a two dimensional Gaussian, with standard deviation $\sqrt{0.1}$, centered at the origin of the system (the results are not sensitive to this choice). The value of p integrated over the domain of computation is conserved and normalized to 1.

To interpret the results of the Fokker-Planck equation presented in section 4, numerical integrations of the SDE are also performed, using the Milstein scheme (Kloeden and Platen 1992) with a time step of 0.1 day. Note that one has to be very careful in choosing a scheme for the accurate numerical integration of a SDE with multiplicative noise (Ewald et al 2004). Even though the Milstein scheme is known to be an accurate scheme, we compared the solutions of the Fokker-Planck equation with the PDFs obtained from numerical integrations of the SDE, and they agreed very well. Although a single

experiment is only one realization of a stochastic process, the ergodic nature of the system ensures agreement with the steady state solution of the corresponding Fokker-Planck equation for the PDF of an infinite number of stochastic realizations.

3. Two Paradigms for Atmospheric Regimes

Obviously, a given non-Gaussian PDF can represent many different dynamical systems. However, it is illustrative to consider two extreme models: a *deterministic model*, in which regimes are entirely due to a nonlinear deterministic \mathbf{A} (perturbed only by state-independent noise), and a *stochastic model*, in which regimes are entirely due to a multiplicative noise term \mathbf{B} (with only a linear \mathbf{A}). The mathematical formulation of the two models is presented in appendix A.

First, we quantify the simple example presented in Fig. 1 by applying these models to the bimodal PDF $p(x) = (1/\sqrt{8\pi}) [\exp(-(x+1.5)^2/2) + \exp(-(x-1.5)^2/2)]$ (see Fig. 2a). Because observed PDFs rarely show any clear multimodality, we also consider the “skewed” PDF given in Fig. 3a, whose departure from Gaussianity is relatively small but is consistent with that observed for weekly averaged circulation anomalies (see section 5). It has a heavier tail than a Gaussian for values $x < -2$ (regime 1), is smaller than a Gaussian for $-2 < x < -0.3$, and is again heavier than a Gaussian for $-0.3 < x < 0.5$ (regime 2). For larger x the PDF is strictly Gaussian.

Given the PDF and $B^2 \equiv 1$, we can solve for $A(x)$ [Eq. (A.3); see appendix A]; results for the bimodal PDF are in Fig. 2b and for the skewed PDF in Fig. 3b. In this case the non-Gaussianity is due to the nonlinearity of the deterministic term $A(x)$. Alternatively, given the PDF and $A(x) = -x$, we can solve for $B(x)$ [Eq. (A.4)]; results for the bimodal PDF are in Fig. 2c and for the skewed PDF in Fig. 3c. Now the non-Gaussianity is due to the structure of the multiplicative noise term $B(x)$. Note (see Fig. 3c) that $B(x)$ for the skewed PDF is approximately piecewise linear, and that as opposed to the stochastic model for the bimodal PDF, the noise amplitude increases for decreasing negative x .

Although the stationary PDFs of the deterministic and stochastic regimes are identical, the conditional PDFs and related mean residence times are not. Here, we define residence time as the time it takes a stochastic trajectory initially at x inside the interval $[x_1, x_2]$ to first leave that interval (see appendix A for more details). For example, the mean residence time in the right peak of the bimodal PDF (interval [1,2]), shown in Fig. 4a, is considerably longer for the deterministic model than for the stochastic model. A similar difference between the two models exists in the interval [-3.5,-2.5] for the weakly skewed PDF (Fig. 5a), but is much less for the interval [-0.25,0.75] (Fig. 5b).

Not surprisingly, the predictability in these two systems is also very different. Here, we define predictability by the expected skill of a perfect model infinite member

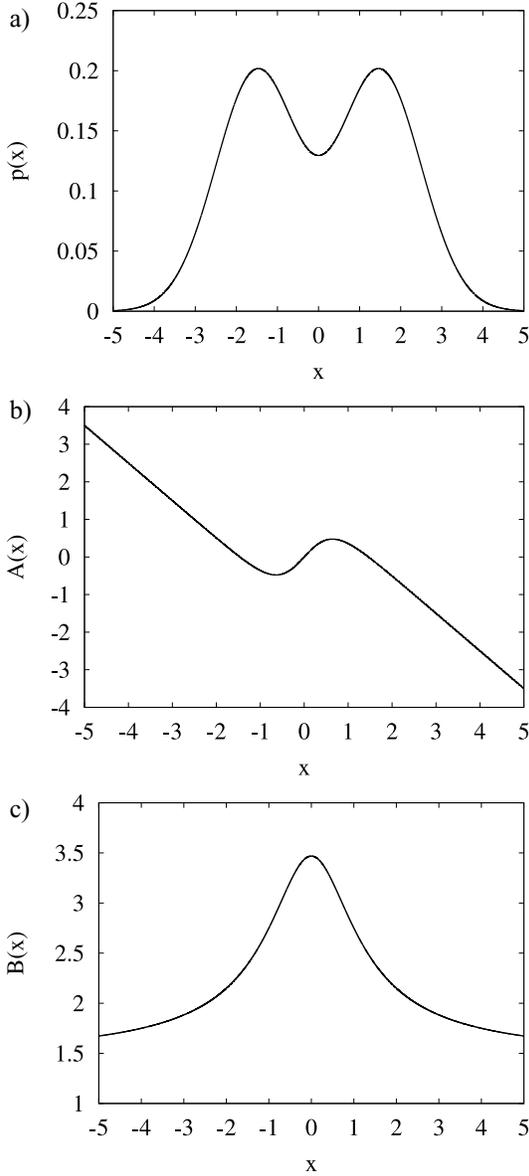


FIG 2. a) Graph of the bimodal PDF $p(x) = (1/\sqrt{8\pi}) [\exp(-(x+1.5)^2/2) + \exp(-(x-1.5)^2/2)]$. b) Solution of $A(x)$ from (A.3), given the bimodal PDF and additive noise $B^2 \equiv 1$. c) Solution of $B(x)$ from (A.4), given the bimodal PDF and a linear deterministic term $A(x) = -x$.

forecast ensemble, measured as an anomaly correlation:

$$\rho_{\infty}(\tau) = \frac{S(\tau)}{\sqrt{S(\tau)^2 + 1}} \quad (6)$$

where $S(\tau)$ is the signal-to-noise ratio and τ is the forecast lead (Sardeshmukh et al 2000; Newman et al 2003). For the bimodal PDF, $\rho_{\infty}(\tau)$ for the initial condition $x_0=1.5$ for the deterministic and the stochastic models is shown in Fig. 4b; $\rho_{\infty}(\tau = 2)$ as a function of initial con-

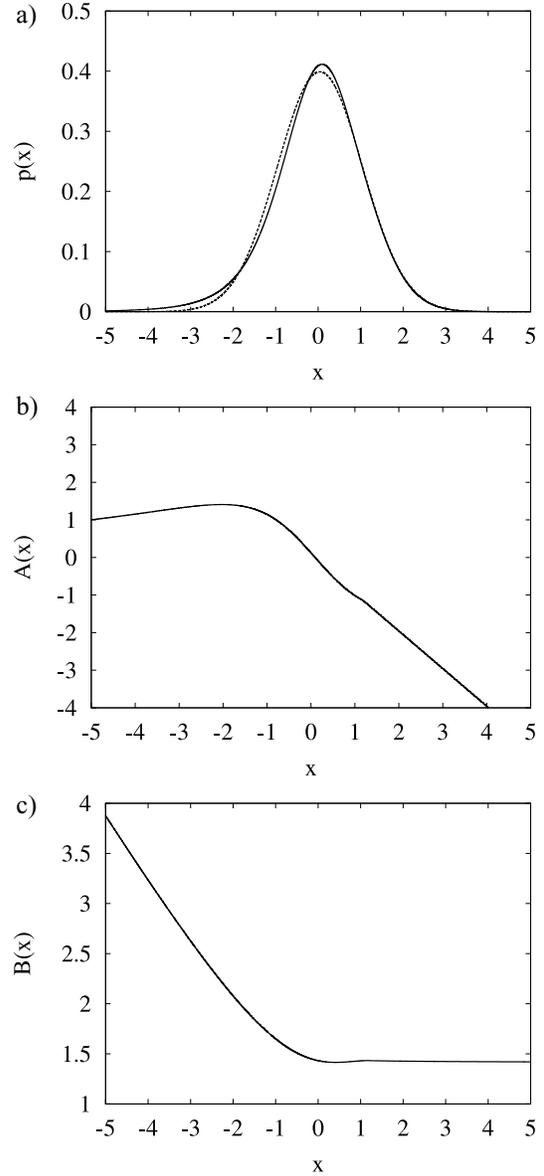


FIG 3. a) Graph of the “skewed” PDF (solid line); the corresponding Gaussian PDF is indicated by the dashed line. b) Solution of $A(x)$ from (A.3), given the bimodal PDF and additive noise $B^2 \equiv 1$. c) Solution of $B(x)$ from (A.4), given the bimodal PDF and a linear deterministic term $A(x) = -x$.

dition is shown in Fig. 4c. As implied by the mean residence times, predictability is much less when the bimodality is due to unpredictable multiplicative noise than when it is due to deterministic nonlinear dynamics. For the skewed PDF, $\rho_{\infty}(\tau)$ for the initial conditions $x_0=-3$, and $x_0=0.25$ are shown in Figs. 5c,d; $\rho_{\infty}(\tau = 2)$ as a function of initial condition is shown in Fig. 5e. Interestingly, predictability for the deterministic model is generally higher for negative x , whereas it is generally higher for positive x in the stochastic model. Thus, a characterization of predictability for a system represented by a given PDF depends on the dynamics of the

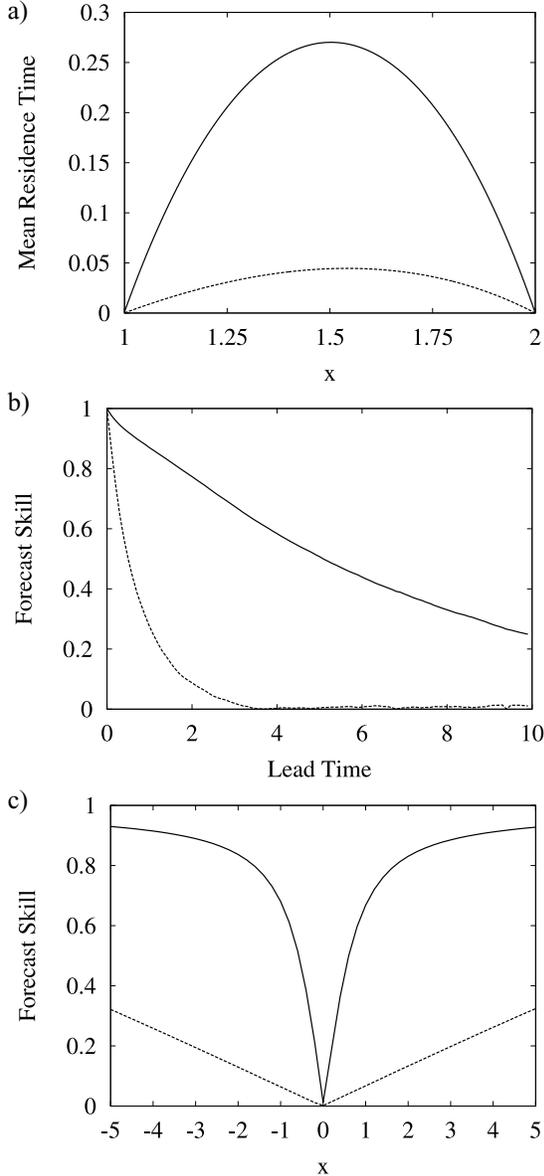


FIG 4. Statistics of the bimodal case (see Fig. 2): a) The mean residence time for the deterministic (solid line) and the stochastic regime (dashed line) in the interval $[1, 2]$. b) ρ_∞ as a function of lead time for the deterministic (solid line) and the stochastic regime (dashed line) for the initial condition $x_0 = 1.5$. c) ρ_∞ as a function of initial condition for the deterministic (solid line) and the stochastic regime (dashed line) for the lead time $\tau = 2$.

underlying system, and cannot be ascertained merely from the existence of non-Gaussianity.

4. Stochastically Perturbed Rossby Waves

In stochastic atmospheric models noise is introduced primarily as an additive process (e.g., Egger 1981; DeSwart 1988; Farrell 1995; Farrell and Ioannou 1996; Newman et al. 1997). However, stochastic forcing may, for example, also represent the fluctuations of model

parameters due to unresolved system components (e.g., Neelin and Weng 1999; Sardeshmukh et al. 2001; Sura 2002; Lin and Neelin 2002). In that case, the stochastic process appears as multiplicative noise, which as is well known can substantially change the dynamical behavior of not only nonlinear systems (Horsthemke and L ef ever 1984; Landa and McClintock 2000; Sura 2002), but also linear systems. Sardeshmukh et al (2001) introduced multiplicative noise to the linearized barotropic vorticity equation, and found that the mean stationary wave response to steady forcing was amplified when the damping parameter fluctuated, but was weakened (in a scale-dependent manner) when the advection parameter fluctuated. Here, we show in the same framework that stochastic damping also results in Rossby waves with a highly non-Gaussian distribution.

4.1 Multiplicative noise in the linearized barotropic vorticity equation

The linearized barotropic vorticity equation is:

$$\frac{\partial}{\partial t} \zeta' = -\nabla \cdot (\bar{\mathbf{v}} \zeta' + \mathbf{v}' \bar{\zeta}) - r \zeta' + F' \quad (7)$$

where ζ is absolute vorticity, \mathbf{v} is the non-divergent horizontal velocity, r is the frictional damping rate, and F is anomalous forcing. Note that F includes both predictable forcing (e.g., steady tropical forcing) and unpredictable forcing, modeled using additive noise (Newman et al 1997). Overbars indicate time means, and primes denote deviations from the time means.

Equation (7) can be written in terms of the streamfunction Ψ ($\zeta' = \nabla^2 \Psi'$) and in operator form as

$$\frac{d\Psi}{dt} = \mathbf{L}\Psi + \mathbf{F} \quad (8)$$

where \mathbf{L} is the linear barotropic operator (e.g., Borges and Sardeshmukh 1995) and (dropping the primes) Ψ and \mathbf{F} are the anomalous streamfunction and forcing vectors, respectively. Any streamfunction anomaly can then be expanded in the basis set of complex eigenvectors \mathbf{E}_j as $\Psi = \sum \Psi_j \mathbf{E}_j$, where $\mathbf{L}\mathbf{E}_j = \lambda_j \mathbf{E}_j$, λ_j are the complex eigenvalues, and Ψ_j are the complex expansion coefficients. The equations for any mode are decoupled from those of all other modes so that for each mode j we can write

$$\frac{d\Psi_j}{dt} = \mathbf{L}_j \Psi_j + \mathbf{f}_j \quad (9)$$

where now the vector Ψ_j represents $(\Psi_{j,r}, \Psi_{j,i})^T$,

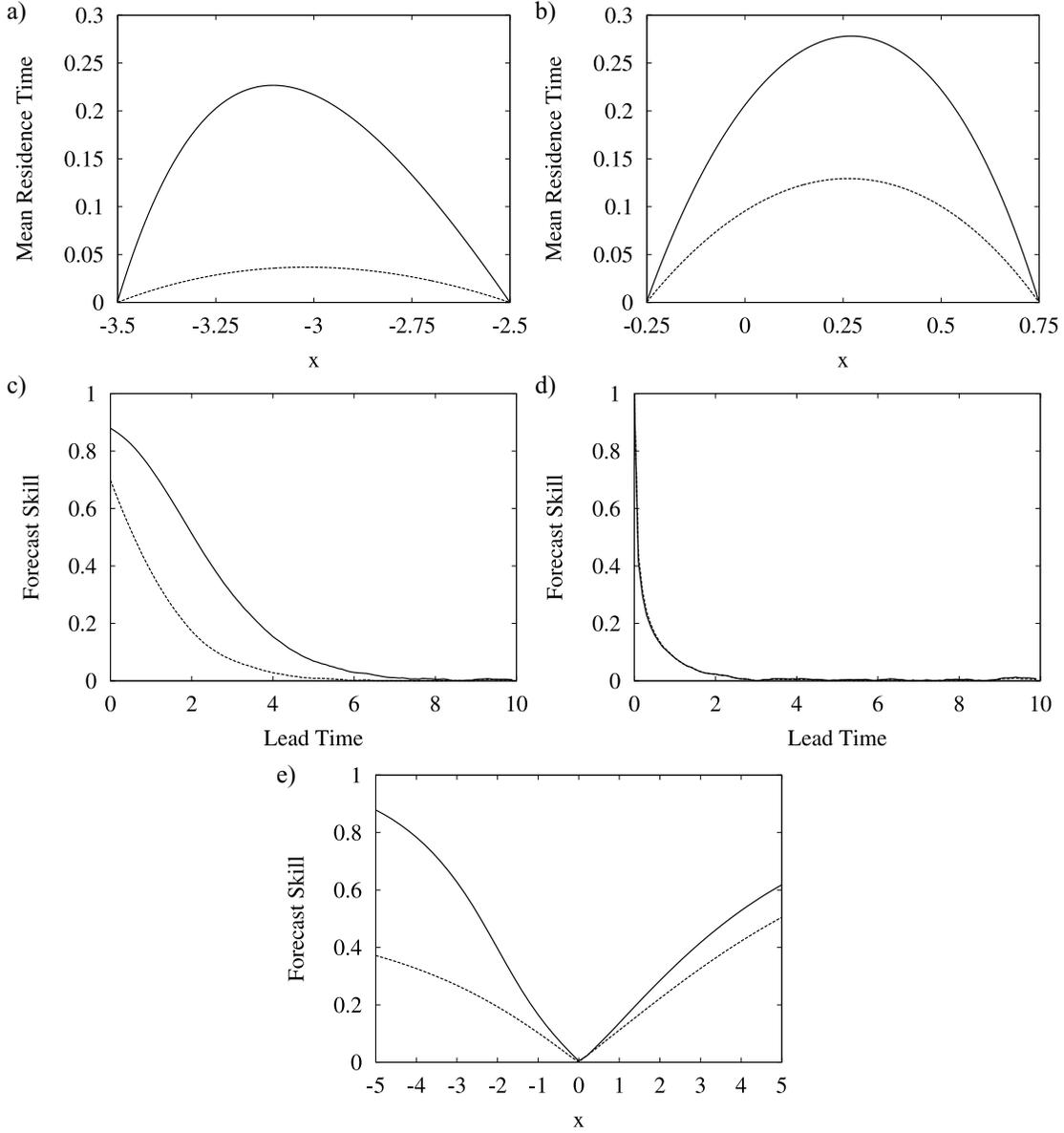


FIG. 5. Statistics of the non-Gaussian case (see Fig. 3): The mean residence times for the deterministic (solid lines) and the stochastic regimes (dashed lines) in the intervals a) $[-3.5, -2.5]$, and b) $[-0.25, 0.75]$. ρ_∞ as a function of lead time for the deterministic (solid line) and the stochastic regime (dashed line) for the initial conditions c) $x_0 = 3$, and d) $x_0 = 0.25$. e) ρ_∞ as a function of initial condition for the deterministic (solid line) and the stochastic regime (dashed line) for the lead time $\tau = 2$.

$$\mathbf{L}_j = \begin{bmatrix} \lambda_{jr} & -\lambda_{ji} \\ \lambda_{ji} & \lambda_{jr} \end{bmatrix} \quad (10)$$

and \mathbf{f}_j represents the projection of \mathbf{F}_j upon the j th mode. For simplicity we now omit the index j .

Because this model does not exhibit variability beyond the decay of damped waves to a steady state, it is assumed that the model can be made more realistic by introducing stochastic perturbations. On the one hand the modes retained in the model can be perturbed by

additive noise. On the other hand, introducing stochastic perturbations in either λ_r or λ_i results in multiplicative noise, and from (9) the Stratonovich-SDE is

$$\frac{d\Psi}{dt} = \mathbf{L}\Psi + \mathbf{f} + \mathbf{B}\eta^M + \eta^A \quad (11)$$

where the forcing is separated into a deterministic component \mathbf{f} and an additive stochastic component η^A . If the decay rate of the wave (λ_r) is perturbed stochastically, then

$$\mathbf{B} = \begin{bmatrix} -\Psi_r & 0 \\ -\Psi_i & 0 \end{bmatrix} \quad (12)$$

If, instead, the phase speed (λ_i) is perturbed stochastically, then

$$\mathbf{B} = \begin{bmatrix} -\Psi_i & 0 \\ \Psi_r & 0 \end{bmatrix} \quad (13)$$

In either case, the multiplicative and additive stochastic noise vectors are $\boldsymbol{\eta}^M = (\eta_r^M, \eta_i^M)$, and $\boldsymbol{\eta}^A = (\eta_r^A, \eta_i^A)$. In the subsequent discussion η_r^M , η_i^M , η_r^A , and η_i^A are assumed to be independent Gaussian white noise processes with corresponding amplitudes σ_r^M , σ_i^M , σ_r^A , and σ_i^A :

$$\begin{aligned} \langle \eta_r^M(t) \rangle &= 0, \quad \langle \eta_r^M(t) \eta_r^M(t') \rangle = (\sigma_r^M)^2 \delta(t-t') \\ \langle \eta_i^M(t) \rangle &= 0, \quad \langle \eta_i^M(t) \eta_i^M(t') \rangle = (\sigma_i^M)^2 \delta(t-t') \\ \langle \eta_r^A(t) \rangle &= 0, \quad \langle \eta_r^A(t) \eta_r^A(t') \rangle = (\sigma_r^A)^2 \delta(t-t') \\ \langle \eta_i^A(t) \rangle &= 0, \quad \langle \eta_i^A(t) \eta_i^A(t') \rangle = (\sigma_i^A)^2 \delta(t-t') \end{aligned} \quad (14)$$

where $\langle \dots \rangle$ denotes the averaging operator. The PDF of this stochastic model is governed by the corresponding Fokker-Planck equation (3) with $\alpha=1/2$.

4.2 Stochastically perturbed Rossby waves on a superrotating basic state

The simplest form of Eq. (7) sets the time mean flow to be in solid body rotation with the earth, where the mean meridional flow is zero and the mean zonal flow is $\bar{u}=u_o \cos(\theta)$. Then the eigenmodes \mathbf{E}_j are just complex spherical harmonics $Y_n^m(\theta, \phi)$ (θ : latitude, ϕ : longitude) for the wavenumber index pair (m, n) (m : zonal wavenumber, n : meridional wavenumber). The Rossby wave dispersion relation gives $\lambda_r = -r$ and $\lambda_i = mD_n$, where

$$\begin{aligned} D_n &= \left[\frac{2}{n(n+1)} + d_n \Delta \right] \Omega \\ d_n &= \frac{2}{n(n+1)} - 1 \end{aligned} \quad (15)$$

Typical parameters of atmospheric motion are used: $r = (4 \text{ d})^{-1}$, $u_o = 15 \text{ ms}^{-1}$, $\Delta = u_o/a\Omega = 0.323$, $\Omega = 2\pi \text{ d}^{-1}$, and mD_n is set to 1 d^{-1} , corresponding to $m=O(5)$ and $n=O(5)$. Steady anomalous forcing is $\mathbf{f}=(1,1)^T$.

Past studies examining the Rossby wave response to steady forcing (e.g., Hoskins et al. 1977; Sardeshmukh

TABLE 1. Skewness and excess kurtosis of the marginal distributions Ψ_r and Ψ_i of the stochastic Rossby wave model with stochastic frictional damping. The excess kurtosis measures the departure from Gaussianity: *excess kurtosis* = *kurtosis* - 3. Note the strong departures (intermittency) from Gaussianity.

Multiplicative noise	Skewness		Excess kurtosis	
	Ψ_r	Ψ_i	Ψ_r	Ψ_i
σ_r^M				
0.0	0.0	0.0	0.0	0.0
0.1	-0.1	0.0	0.2	0.2
0.2	-0.3	-0.1	1.4	1.3
0.3	-0.7	-0.3	6.6	7.5
0.4	-2.7	-1.4	99.7	83.7

and Hoskins 1988) have typically considered r and u_o to be fixed. However, a more realistic representation of Rossby wave propagation on the sphere might also consider stochastic perturbations in r and u_o (Sardeshmukh et al. 2001). In the following, we solve the Fokker-Planck equation (3) for p using these parameters and stochastically perturbing either r or u_o for different values of the noise amplitude σ_r^M . Weak additive noise with amplitudes $\sigma_r^A = \sigma_i^A = 0.2$ is also included. Note that due to the structure of the matrices (12) and (13) the imaginary part of the multiplicative noise η_i^M (with amplitude σ_i^M) has no impact and is, therefore, not specified. In both experiments the strength of the multiplicative noise σ_r^M is scaled by its effect on the variance of the system. The multiplicative noise amplitude is increased until the variances $\langle \Psi_r^2 \rangle$ and $\langle \Psi_i^2 \rangle$ reach approximately 1. This yields the interval $\sigma_r^M = 0.0, 0.1, \dots, 0.4$ for the stochastic damping, and the interval $\sigma_r^M = 0.0, 0.2, \dots, 0.8$ for the stochastic basic state.

Stochastic damping

The marginal PDFs for varying noise amplitude in the damping are shown in Figs. 6a,b. Table 1 shows the values of the skewness and the excess kurtosis (the excess kurtosis measures the departure from Gaussianity: *excess kurtosis*=*kurtosis*-3). If the multiplicative noise is nonexistent or weak ($\sigma_r^M = 0.0, 0.1, 0.2$) the marginal distributions $p(\Psi_r)$ and $p(\Psi_i)$ are approximately Gaussian. However, for stronger noise amplitudes ($\sigma_r^M = 0.3, 0.4$) the distributions become skewed with heavy, highly non-Gaussian tails. A sample time series of Ψ_i of the stochastic Rossby wave model with relatively strong stochastic frictional damping ($\sigma_r^M = 0.4$) is shown in Fig. 7a. The distinct feature of the time series is its intermittent behavior. It is this intermittency that gives rise to the heavy, non-Gaussian tails of the PDFs, meaning that extreme events are far more probable than can be expected from a Gaussian distribution. The positive excess kurtosis is a measure of this intermittency.

Stochastic basic state

Unlike the stochastic damping case, increasing noise

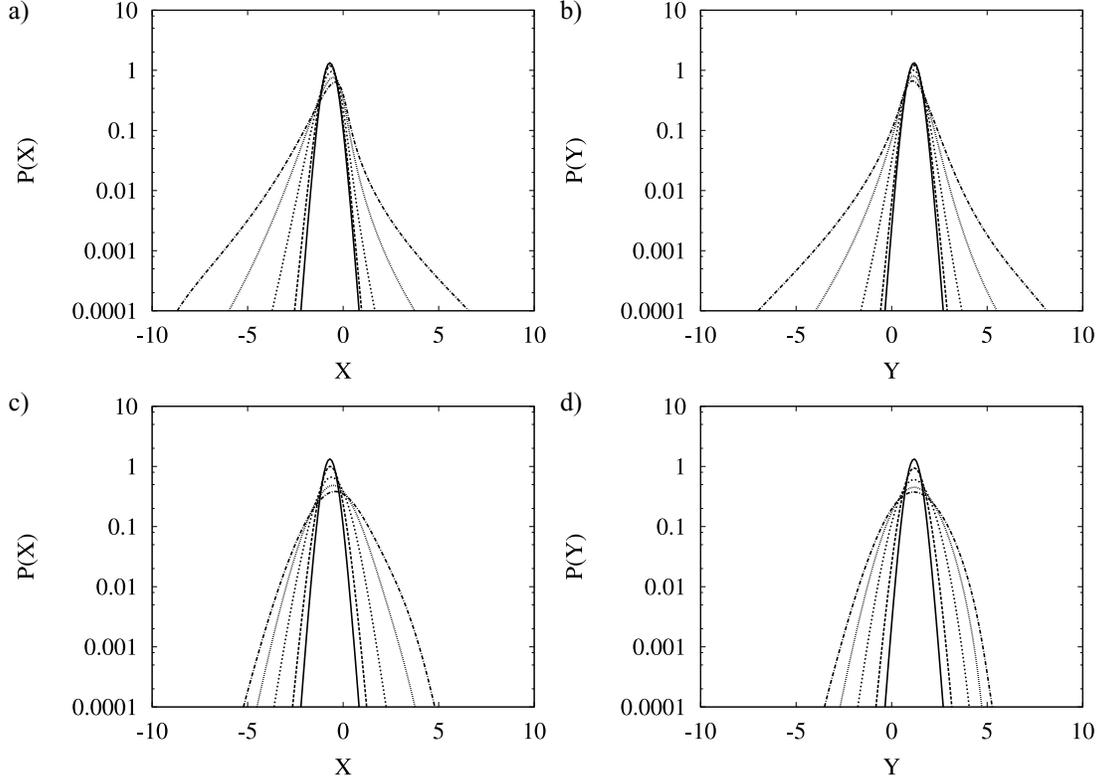


FIG 6. a-b) Steady state marginal PDF of the Rossby wave model with stochastic frictional damping for $\sigma_r^M = 0.0$ (solid line), 0.1 (long dashed line), 0.2 (short dashed line), 0.3 (dotted line), and 0.4 (dotted-dashed line): a) $p(\Psi_r)$; X denotes Ψ_r , b) $p(\Psi_i)$; Y denotes Ψ_i . c-d) Steady state marginal PDF of the Rossby wave model with stochastic basic state for $\sigma_r^M = 0.0$ (solid line), 0.2 (long dashed line), 0.4 (short dashed line), 0.6 (dotted line), and 0.8 (dotted-dashed line): c) $p(\Psi_r)$; X denotes Ψ_r , d) $p(\Psi_i)$; Y denotes Ψ_i .

in the superrotation does not produce heavy-tailed distributions, as is evident in both the marginal PDFs (Figs. 6c,d) and the skewness and excess kurtosis in Table 2. A sample time series of Ψ_i with a relatively strong stochastic basic state ($\sigma_r^M=0.8$) (Fig. 7b) does not show the intermittent behavior seen in Fig. 7a. Because the stochasticity in the basic state only influences the phase of the Rossby wave, the PDF of the system remains more nearly Gaussian, while multiplicative noise in the frictional damping alters the energy of the wave and causes the PDF to be highly non-Gaussian (see appendix B and Fig. 12 for more details).

TABLE 2. Skewness and excess kurtosis of the marginal distributions Ψ_r and Ψ_i of the stochastic Rossby wave model with a stochastic base state. The excess kurtosis measures the departure from Gaussianity: *excess kurtosis* = *kurtosis* - 3. Note that the distributions are nearly Gaussian, even for strong multiplicative noise.

Multiplicative noise σ_r^M	Skewness		Excess kurtosis	
	Ψ_r	Ψ_i	Ψ_r	Ψ_i
0.0	0.0	0.0	0.0	0.0
0.1	0.0	0.0	0.0	0.2
0.2	0.0	-0.1	0.1	0.0
0.3	-0.1	-0.2	0.3	0.2
0.4	-0.2	-0.4	0.9	0.9

4.3 Departures from bivariate Gaussianity

As can be expected from the marginal PDFs, stochastic damping produces notable departures from bivariate Gaussianity. This is true not only for parameters representing Rossby waves on a superrotating flow, but also for parameters representing the more general problem of

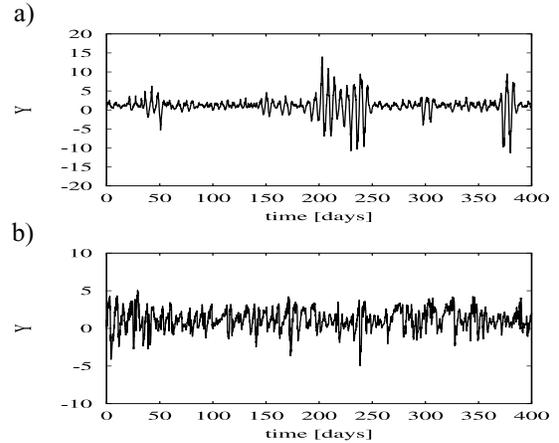


FIG. 7. Sample time series of Ψ_i of the stochastic Rossby wave model with a) stochastic frictional damping ($\sigma_r^M = 0.4$) and b) stochastic basic state ($\sigma_r^M = 0.8$). Y denotes Ψ_i . Note the different scales on the ordinates.

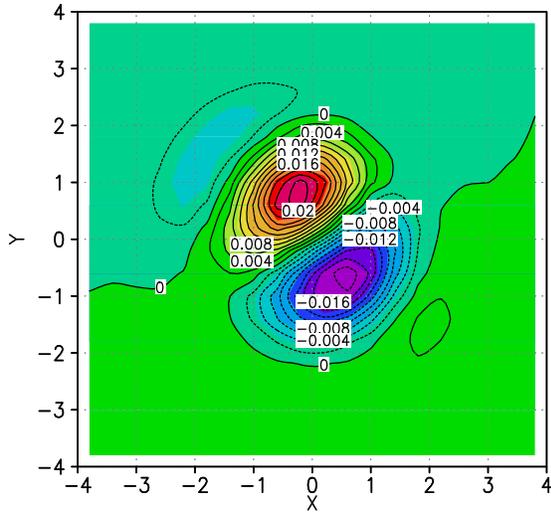


FIG 8. Steady state PDF anomalies of the Rossby wave model ($\lambda_r = 0.07$ and $\lambda_i = 0.19$) with a stochastic frictional damping ($\sigma_r^M = 0.05$). The axes are rotated so that $X \equiv -0.86\Psi_r - 0.51\Psi_i$ and $Y \equiv 0.51\Psi_r - 0.86\Psi_i$. The contour interval is 0.002.

large scale barotropic Rossby waves evolving on a zonally and meridionally varying base state. As an example, we set $\lambda_r = 0.07$ and $\lambda_i = 0.19$ in (10), corresponding to a period of about 33 days and an e-folding time of 14 days, typical of the least damped eigenmode of the 250 hPa climatological December--February flow (Borges and Sardeshmukh 1995). We set $\sigma_r^M = 0.05$, and for simplicity set the additive noise to zero; including additive noise tends to decrease skewness, which can be offset by increasing F , but has much less effect upon the heavy tails. For display purposes, we remove the mean drift, rotate the PDF so that the abscissa points in the direction of maximum covariance, normalize by the variance along both axes, and show the difference between the PDF and the corresponding bivariate Gaussian distribution.

The resulting departure from Gaussianity (Fig. 8) appears as two slightly arcing ridge/trough (positive/negative departures) pairs aligned roughly along a line. This pattern is qualitatively similar for a wide parameter range. Decreasing either λ_r or λ_i strengthens the departures from Gaussianity, suggesting that the lower frequency and least damped barotropic eigenmodes could have the greatest non-Gaussian behavior. Changing these parameters also can change the covariance between the real and imaginary parts of Ψ and thus the rotation applied in Fig. 8.

5. Observed Atmospheric PDFs

Having demonstrated that non-Gaussian behavior can result from simple linear multiplicative noise, we next compare the observed non-Gaussianity to the multiplicative stochastic paradigm. Following previous studies (e.g., Mo and Ghil 1988; Molteni et al. 1990; Kimoto

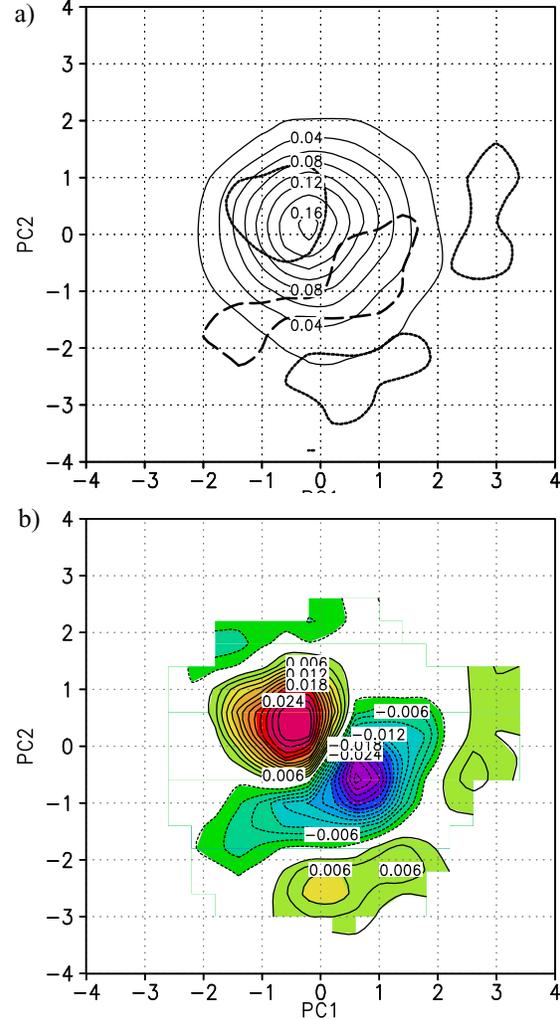


FIG. 9. a) PDF of PC1 and PC2 (thin solid lines). Regions of significant positive deviations from a bivariate Gaussian PDF are indicated by thick solid lines, significant regions of negative deviations are indicated by thick dashed lines. The deviations from a bivariate Gaussian PDF are significant at the 90% confidence level. b) The actual PDF anomalies of PC1 and PC2. The zero contour is omitted for clarity, the contour interval is 0.002.

and Ghil 1993a,b; Corti et al. 1999; Smyth et al. 1999; Weisheimer et al. 2001), we consider non-Gaussian regimes in a highly truncated bivariate phase space spanned by two leading EOFs.

5.1 Bivariate PDF

The analysis is applied to Northern Hemisphere 750 hPa streamfunction data for the extended winters (November--March) of 1949/50--2001/02, spectrally truncated to T21 resolution. The data were obtained from the National Centers for Environmental Prediction (NCEP) Reanalysis dataset (Kalnay et al. 1996). Streamfunction anomalies were defined by removing the seasonal cycle (that is, the annual mean plus the first three

annual harmonics) from each variable at each gridpoint and then applying a 7-day running mean filter. Calculations using 500 hPa geopotential height pentad data yielded similar results (not shown).

A Principal Component Analysis (PCA) was applied to the streamfunction anomalies. The first two EOF patterns form the orthogonal basis vectors of a reduced phase space. The first EOF (EOF1 with principal component PC1) explains 16.6% of the total variance, whereas the second EOF (EOF2 with principal component PC2) explains 8.2%. The EOFs (not shown) resemble those found in earlier studies (e.g., Kimoto and Ghil 1993a; Smyth et al. 1999); signs are defined as in Kimoto and Ghil (1993a). The PCs are normalized to have zero mean and unit standard deviation. Monahan et al. (2003) suggested that the leading atmospheric regimes have a large projection on PCs 1 and 2 and a small, but non-negligible, projection on PC3. Including this projection on PC3 did not affect any of the results shown below.

The bivariate PDF of PC1 and PC2, determined by dividing the interval $[-4:4, -4:4]$ into 20×20 equal bins and then applying a 3×3 bin smoothing, is shown in Fig. 9a. Although the PDF is generated with a fairly simple technique, virtually identical results are obtained using bivariate Gaussian kernel density estimation with bandwidth $h=0.2$ (not shown). Positive (negative) departures from Gaussianity (Fig. 9b) indicate that the observed PDF is greater (smaller) than the corresponding bivariate Gaussian distribution.

Past studies of bivariate PDFs such as shown here tend to focus on the local maxima, determined for example using some bump-hunting algorithm, identifying them as regimes and producing corresponding composite anomaly maps. Such maxima can, however, be sensitive to the dataset used, to sampling considerations (particularly the data period), and to the smoothing applied to the PDF. Considerably more robust is the large scale pattern of the departures from bivariate Gaussianity, consisting of two slightly arcing ridge/trough pairs aligned roughly along a line from the upper left quadrant to the lower right quadrant. Note in particular the pronounced negative region; although this feature has amplitude as large as the positive departures, it has generally received far less attention. This pattern exists in many different subsamples of the data: using only the odd years, only the even years, only the years 1949–1975, or only the years 1976–2001. Also, broad regions of positive and negative departures are statistically significant at the 90% confidence level (heavy contours in Fig. 9a) determined using the Monte-Carlo method employed in Kimoto and Ghil (1993a).

5.2 Balancing the probability budget

Using the Fokker-Planck equation (3), a steady climatological probability budget for a bivariate system ($i=1,2; j=1,2$), assumed to be Markovian, is

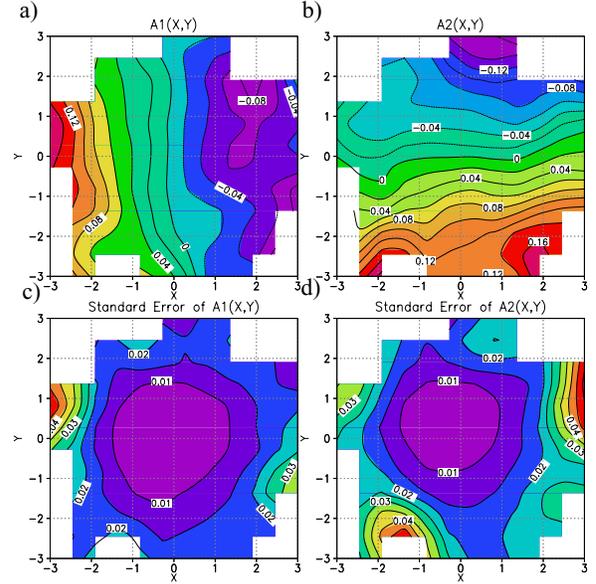


FIG. 10. The effective drift (and its standard error) estimated from data. a) $A1(x, y)$ denotes the x-component (PC1), and b) $A2(x, y)$ denotes the y-component (PC2) of the two-dimensional system. The contour and shading interval is 0.02. c) and d) are the related standard errors with a contour and shading interval 0.005.

$$-\sum_i \frac{\partial}{\partial x_i} A_i p(\mathbf{x}, t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\mathbf{B} \mathbf{B}^T)_{ij} p(\mathbf{x}, t) = 0 \quad (16)$$

That is, in a time-averaged sense, the effective drift balances the stochastic diffusion. We can then use Eq. (16) to ask whether the nonlinear effective drift due to interaction between PC1 and PC2, estimated by using its finite difference approximation (4), is sufficient to produce the non-Gaussianity of the observed PDF. That is, we ask whether or not the probability budget can be balanced using purely additive stochastic forcing.

The effective drift (and the related standard errors) of the data components are shown in Fig. 10. The uncertainties are relatively small: the standard error is about 0.01–0.02 for most of the points. Unfortunately, the errors in a direct estimation of the diffusion term are much larger (Sura and Barsugli 2002), and so we use a different tack seeing whether additive noise can balance the probability budget. First, we rewrite Eq. (16) for the special case where \mathbf{B} is purely additive ($i=1,2; j=1,2$):

$$-\sum_i \frac{\partial}{\partial x_i} A_i p(\mathbf{x}, t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\mathbf{B}_A \mathbf{B}_A^T)_{ij} p(\mathbf{x}, t) = R \quad (17)$$

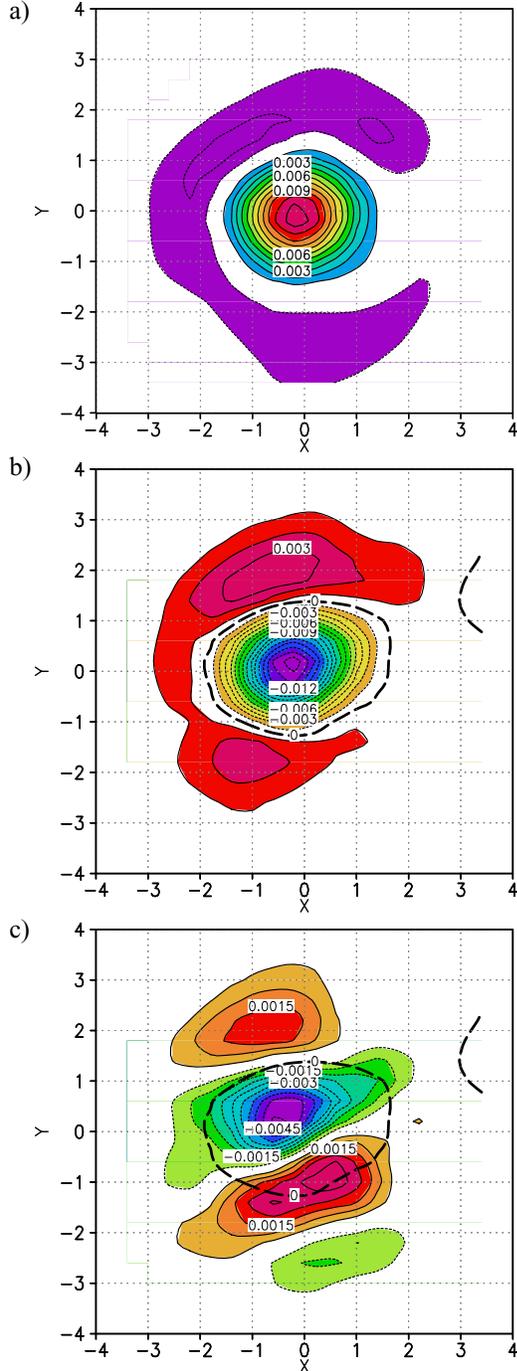


FIG. 11. a) The effective deterministic term (C.I.= 0.001), b) the stochastic term (C.I. = 0.001, the dark dashed line indicates the zero contour; see text for details), and c) the sum of the effective deterministic and the stochastic term of the steady Fokker-Planck equation (17) evaluated by using the PDF (Fig. 9), the deterministic drift (Fig. 10), and pure additive noise given by the matrix (19); C.I. = 0.0005, the dark dashed line indicates the zero contour of the stochastic term; see text for details. The negative value of that sum ($-R$) can be interpreted as the multiplicative noise contribution of the Fokker-Planck equation (17); positive (negative) values mean that the additive noise is too weak (strong) to balance the deterministic term. The shown regions are significant at the 90% confidence level. The x-component represents PC1, the y-component represents PC2.

From Eq. (17), it is clear that zeroes in the diffusion's contribution to the budget are due solely to the inflexion points of the PDF $p(\mathbf{x}, t)$. Therefore, given the empirically-determined PDF, a map of this term with any arbitrarily chosen additive noise will have the same zero lines as the map corresponding to any additive noise, although the maps themselves might change greatly. Of course, if the probability budget balances, the zero lines of the diffusion's contribution would be the zero lines of the drift's contribution as well.

Although any additive noise will do, it is most satisfying to use a noise matrix with some approximation to truth using a linear approximation to $\mathbf{A}(\mathbf{x})$. We emphasize that the only time this approximation is used is in estimating a reasonable test matrix for the noise properties. Still, the approximation is not bad, and a linear fit to $\mathbf{A}(\mathbf{x})$, $\bar{\mathbf{A}} \mathbf{x}$, corresponds very well to the parameters obtained from a linear inverse modeling procedure (e.g., Penland 1989; Penland and Sardeshmukh 1995; Winkler et al. 2001). The linear drift matrix $\bar{\mathbf{A}}$ is

$$\bar{\mathbf{A}} = \begin{bmatrix} -0.050 & -0.008 \\ 0.007 & -0.081 \end{bmatrix} \quad (18)$$

The corresponding additive noise matrix $\mathbf{B}_A \mathbf{B}_A^T$ is then determined from the Fluctuation-Dissipation relation (e.g., Penland and Matrosova 1994; Penland and Sardeshmukh 1995):

$$\mathbf{B}_A \mathbf{B}_A^T = \begin{bmatrix} 0.100 & 0.001 \\ 0.001 & 0.161 \end{bmatrix}. \quad (19)$$

Using the PDF (Fig. 9), the nonlinear effective drift $\mathbf{A}(\mathbf{x})$ (Fig. 10), and pure additive noise given by the matrix (19), the two terms on the left hand side of the Fokker-Planck equation (17) are evaluated and shown in Figs. 11a and 11b. The only important aspect of Fig. 11b is the dark dashed line indicating the zero contour. The sum (the residual R) (significant at the 90% confidence level, see below) of both terms is shown in Fig. 11c. Note that the zero contour in Fig. 11b, reproduced in Fig. 11c, crosses significantly large values of the residual, indicating that multiplicative noise is necessary to balance the budget.

If we partition the net noise effect $\mathbf{B} \mathbf{B}^T$ in (17) into an additive noise term $\mathbf{B}_A \mathbf{B}_A^T$ and a multiplicative noise term $\mathbf{B}_M \mathbf{B}_M^T$, then $-R$ would represent the multiplicative part:

$$R = -\frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\mathbf{B}_M \mathbf{B}_M^T)_{ij} p(\mathbf{x}, t) \quad (20)$$

with R and $\mathbf{B}_A \mathbf{B}_A^T$ appropriately scaled so that the appropriate Fluctuation-Dissipation relation is satisfied.

In summarizing, if the departures from Gaussianity are

primarily due to the nonlinear drift term, then the probability budget for the joint PC1/2 PDF would be balanced with pure additive noise. Clearly, it is not; all three terms of the Fokker-Planck equation have the same order of magnitude. Positive (negative) values mean that the additive noise is too weak (strong) to balance the drift. Thus, the observed departures from Gaussianity do not result from the nonlinear drift term defined in the two-dimensional space, but rather from the multiplicative structure of the noise.

Despite the relative shortness of the data record, this result appears robust. For example, it is virtually unchanged if only the data subsamples listed above are used. Moreover, the residual is statistically significant, determined as before using a Monte-Carlo method. One hundred time series having the same covariances and lag-one autocorrelations as the original time series were generated, and the deterministic and additive stochastic terms (and their sum) in the Fokker-Planck equation were computed for each of them. The number of those random terms larger/smaller than the original terms were used to obtain confidence intervals at each point in the bivariate phase space.

6. Summary and Conclusions

In this paper we outlined a stochastic perspective on atmospheric regime behavior based on a treatment of climate variability as a stochastic system with state dependent noise. We have demonstrated how some simple linear (or nearly linear) systems with multiplicative noise can produce non-Gaussian regime-like behavior without multiple equilibrium solutions of the governing equations. The presence of non-Gaussianity, therefore, does not by itself imply that a system has deterministic nonlinear multiple regimes, nor that these regimes have a noticeable enhancement of persistence or predictability due to this nonlinearity.

The linear Rossby wave response to steady forcing is non-Gaussian when the damping contains a stochastic component, as might be expected, for example, from including gustiness in a linear parameterization of boundary layer dissipation. This multiplicative noise leads to a non-Gaussian distribution due to an intermittent behavior of the Rossby waves. Many nonlinear systems are intermittent in time, space, or both (e.g., Sreenivasan 1999; Sreenivasan and Antonia 1997). Again, the most common explanation for intermittency is that they are induced by nonlinearities in the slow manifold of the governing equations.

The similarity between Figs. 8 and 9b, and the analysis of the probability budget of the bivariate PC1/2 PDF, suggests that atmospheric variability may be represented by a phase space which can be characterized not so much by localized regimes as by less-localized departures from Gaussianity that are largely a result of an effectively almost linear system perturbed by multiplicative noise. This is also consistent with results from

linear inverse modeling (LIM; Winkler et al 2001; Newman et al 2003), in which observed weekly variability of 250 hPa and 750 hPa streamfunction is successfully modeled and predicted, even in geographic regions where non-Gaussianity is relatively large, with a multivariate linear system (with $O(30)$ degrees of freedom) plus noise. In fact, LIM cannot distinguish between a linear system driven by additive noise (which also need not be Gaussian), versus one that also includes linear multiplicative noise. In the former case, LIM returns the deterministic linear operator \mathbf{L} , and the “best” forecasts of \mathbf{x} at forecast lead τ are $\mathbf{x}(t+\tau) = \mathbf{G}(\tau)\mathbf{x}(t) = \exp(\mathbf{L}\tau)\mathbf{x}(t)$. In the latter case LIM returns the effective linear operator $\tilde{\mathbf{L}} = \mathbf{L} + (1/2)\mathbf{B}^2$ (that is, the noise-induced drift is part of the linear operator returned by LIM), and the “best” forecasts are $\mathbf{x}(t+\tau) = \mathbf{G}(\tau)\mathbf{x}(t) = \exp(\tilde{\mathbf{L}}\tau)\mathbf{x}(t)$. Trying to differentiate between these two models is a major focus of our future research. In light of our current results, success of the LIM suggests that observed non-Gaussianity is due to the structure of the noise, that is, to the fast component of the nonlinear system. In this view, the most predictable states of the atmosphere, which correspond to the growing singular vectors of \mathbf{G} , will be affected by how this fast nonlinearity affects \mathbf{G} through \mathbf{B} .

However, while our observational results are suggestive, they do not constitute a proof. We do not yet have a solution for what $\mathbf{B}(\mathbf{x})$ should be for the real atmosphere, either by determining it from the extremely limited dataset, from extended model datasets, or from a simplification of the equations of motion. Furthermore, because of the highly truncated space used, while the probability budget is consistent with multiplicative noise forcing as the cause of the observed non-Gaussian bivariate PC1/2 PDF, it may also be consistent with slow nonlinear interactions of PC1 and/or PC2 with higher-order EOFs. In fact, what may be most clear from these results is that trying to understand -- or even define -- observed atmospheric regime behavior in a highly truncated phase space, as has generally been the approach of virtually all past studies, has the potential to be highly misleading.

Finally, we note that these two paradigms of non-Gaussian atmospheric behavior do not have to be mutually exclusive. In reality, there may not be such a clean separation between the two since there may not be such a clean separation of timescales in the atmosphere. Nevertheless, a comprehensive approach towards a better understanding of atmospheric regime behavior must consider state-dependent noise. Further research is required to assess the more detailed extent to which multiple non-Gaussian regimes may be due to the nature of unpredictable stochastic forcing, rather than to the slow, predictable deterministic nonlinear dynamics of the atmosphere.

Acknowledgements. This work was partially funded by the Predictability DRI of the Office of Naval Research, grant N00014-99-1-0021.

APPENDIX A

Fitting a PDF to deterministic and stochastic models

In this appendix we show how various combinations of different deterministic and stochastic terms in the governing SDE can give rise to a given PDF. We consider a univariate Itô SDE of the form

$$dx = A(x)dt + B(x)dW \quad (\text{A.1})$$

where $A(x)$ and $B(x)$ are sufficiently smooth and bounded functions, and W denotes a Wiener process. The PDF $p(x, t)$ of the Itô SDE (A.1) is governed by the corresponding Itô-Fokker-Planck equation (e.g., Gardiner 1985; Horsthemke and L  f  ver 1984; Paul and Baschnagel 1999):

$$\frac{\partial}{\partial t}p(x, t) = -\frac{\partial}{\partial x}A(x)p(x, t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}B(x)^2p(x, t) \quad (\text{A.2})$$

Given a stationary non-Gaussian PDF $p(x)$ (with $p(-\infty) = p(\infty) = 0$) and pure additive noise $B^2 \equiv \text{const}$, the stationary Fokker-Planck equation can be solved for $A(x)$:

$$A(x) = \frac{B^2}{2}\frac{\partial}{\partial x}\ln p(x) + \frac{\beta}{p(x)} \quad (\text{A.3})$$

with the arbitrary constant β . Given the same PDF $p(x)$ and the deterministic damping term $A(x)$, the stationary Fokker-Planck equation can be solved for $B(x)$:

$$B(x) = \left(\frac{-2}{p(x)} \int_{-\infty}^x [A(x')p(x') + \beta] dx' \right)^{1/2} \quad (\text{A.4})$$

with an arbitrary constant β . That means that the same PDF can either be produced by nonlinear deterministic dynamics with additive noise, or, if $A(x) \propto -x$, by linear deterministic dynamics with multiplicative noise. Because we expect $A(x)$ and $B(x)$ to be bounded in physically reasonable situations, the constant is set to zero: $\beta=0$.

Even if deterministic and stochastic regimes do have the same stationary PDFs, the dynamical properties (and, therefore, the predictabilities) of the trajectories are fundamentally different. This behavior can be illustrated by the mean residence times. The residence time of a stochastic trajectory initially at x inside the interval $[x_1, x_2]$ is the time until the trajectory first hits the boundary of the interval. The mean residence time $\tau(x)$ of a stochastic trajectory governed by the SDE (A.1) is given by (e.g., Gardiner 1985; Horsthemke and L  f  ver 1984; Paul and Baschnagel 1999)

$$A(x)\frac{\partial \tau}{\partial x} + \frac{1}{2}B(x)^2\frac{\partial^2 \tau}{\partial x^2} + 1 = 0 \quad (\text{A.5})$$

with the boundary conditions $\tau(x_1) = \tau(x_2) = 0$.

APPENDIX B

An analytical solution of a Fokker-Planck equation with multiplicative noise

In this appendix an analytical solution of the Itô and Stratonovich Fokker Planck equation for a stochastic basic state and no steady Rossby wave forcing [$\mathbf{S} = (0, 0)^T$] is presented. This solution is then used to discuss the fundamental physical difference between a stochastic frictional damping and a stochastic basic state. The SDE for the time evolution of the vector $\Psi = (\Psi_r, \Psi_i)^T$ is

$$\frac{d\Psi}{dt} = \mathbf{L}\Psi + \mathbf{B}\eta^M + \eta^A \quad (\text{B.1})$$

with the matrices

$$\mathbf{L} = \begin{bmatrix} -r_o & -mD_n \\ mD_n & -r_o \end{bmatrix} \quad (\text{B.2})$$

and

$$\mathbf{B} = \begin{bmatrix} -\Psi_i & 0 \\ \Psi_r & 0 \end{bmatrix} \quad (\text{B.3})$$

and the multiplicative and additive stochastic noise vectors are $\eta^M = (\eta_r^M, \eta_i^M)$, and $\eta^A = (\eta_r^A, \eta_i^A)$. The stochastic components η_r^M , η_i^M , η_r^A , and η_i^A are assumed to be independent Gaussian white noise processes with corresponding amplitudes σ_r^M , σ_i^M , σ_r^A , and σ_i^A [Eq. 14]. Note that due to the structure of the matrix (B.3) the imaginary part of the multiplicative noise η_i^M (with amplitude σ_i^M) has no impact and is, therefore, not specified. For identical additive noise amplitudes $\sigma_r^A = \sigma_i^A = \sigma^A$ the corresponding Fokker-Planck equation for the PDF $p(\Psi, t)$ reads ($i = 1, 2; j = 1, 2$):

$$\begin{aligned} \frac{\partial}{\partial t}p(\Psi, t) = & \quad (\text{B.4}) \\ & -\sum_i \frac{\partial}{\partial \Psi_i} \left[(\mathbf{A}\Psi)_i + \alpha \sum_{j,k} (\sigma_r^M)^2 \left(\frac{\partial B_{ik}}{\partial \Psi_j} \right) B_{jk} \right] p(\Psi, t) \\ & + \frac{1}{2} \sum_{i,j} (\sigma_r^M)^2 \frac{\partial^2}{\partial \Psi_i \partial \Psi_j} (\mathbf{B}\mathbf{B}^T)_{ij} p(\Psi, t) \\ & + \frac{1}{2} \sum_i (\sigma^A)^2 \frac{\partial^2}{\partial \Psi_i^2} p(\Psi, t) \end{aligned}$$

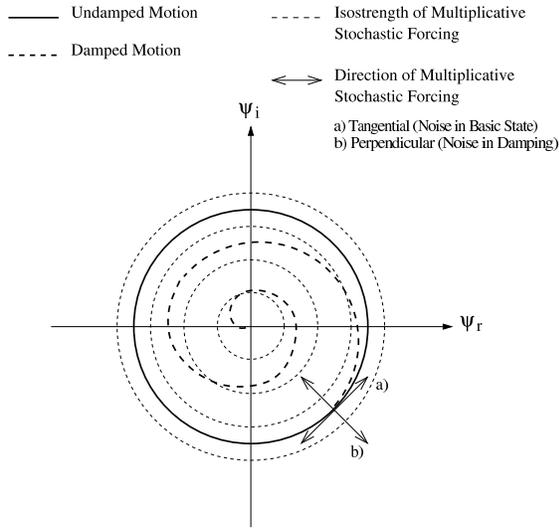


FIG. 12. A schematic sketch to show the fundamental physical difference between a stochastic frictional damping and a stochastic basic state. The thick solid circle visualizes the energy conserving wave motion of the undamped system. The thick dashed line shows the phase space motion of the corresponding damped system. The thin dashed circles are isolines of the strength [defined as $(\sigma_r^M \Psi_r)^2 + (\sigma_i^M \Psi_i)^2$] of the multiplicative stochastic forcing. The arrows show the directions of the stochastic kicks in a) the zonal basic state (tangential) and b) the frictional damping (perpendicular). See appendix B for a more detailed discussion.

with $\alpha=0$ for Itô systems and $\alpha=1/2$ for Stratonovich systems. Because of the identical additive noise amplitudes the Fokker-Planck equation has an exact stationary Gaussian solution even in the presence of multiplicative noise with amplitude σ_r^M :

$$p(\Psi) = \Phi \exp(-\gamma(\Psi_r^2 + \Psi_i^2)) \quad (\text{B.5})$$

with the normalization constant Φ , and

$$\gamma = \left(\frac{r_o - \frac{1}{2}(\sigma_r^M)^2 + \alpha(\sigma_r^M)^2}{(\sigma^A)^2} \right). \quad (\text{B.6})$$

Note that the noise-induced drift in the Stratonovich system ($\alpha=1/2$) effectively increases the damping of the mean flow. Yet, the noise-induced drift is compensated by the remaining effect of the multiplicative noise. Thus, in the Stratonovich calculus there is no effect of the multiplicative noise on the stationary PDF at all. This behavior can be understood by simple geometric considerations described below. These considerations also clarify the fundamental physical difference between a stochastic frictional damping and a stochastic basic state. A schematic explanation is shown in Fig. 12. The equation of a circle (the undamped motion of the Rossby wave with or without the presence of additive noise with

identical amplitudes $\sigma_r^A = \sigma_i^A = \sigma^A$) with radius a around the origin of the coordinate system in the (Ψ_r, Ψ_i) -plane is

$$f(\Psi_r, \Psi_i) = \Psi_r^2 + \Psi_i^2 = a^2 \quad (\text{B.7})$$

A vector perpendicular to the circle (B.7) is $\nabla f(\Psi_r, \Psi_i) \propto (-\Psi_r, -\Psi_i)$. A vector tangential to the circle (B.7) is $\nabla f(\Psi_r, \Psi_i) \propto \mathbf{e}_z \propto (-\Psi_i, \Psi_r)$, where \mathbf{e}_z is the unit vector perpendicular to the (Ψ_r, Ψ_i) -plane. It is clear from Eqs. 12 and 13, and the multiplicative noise vector, that the kicks of the stochastic damping are perpendicular to the undamped circular phase space motion, whereas the stochastic kicks in the basic state are tangential to it. Therefore, the stochasticity in the zonal basic state u_0 only influences the phase of the Rossby wave, but not its energy.

When stochastic perturbations occur in the frictional damping r_0 , the energy is altered, but not the phase of the Rossby wave. In Fig. 12 the thick solid circle indicates the energy conserving wave motion of the undamped system with or without the presence of additive noise with identical amplitudes. The thick dashed line shows the phase space motion of the corresponding damped system. The thin dashed circles are isolines of the strength [defined as $(\sigma_r^M \Psi_r)^2 + (\sigma_i^M \Psi_i)^2$] of the multiplicative stochastic forcing. The arrows show the directions of the stochastic kicks in a) the zonal basic state and b) the frictional damping.

Because the stochasticity in the basic state only influences the phase of the Rossby wave, the PDF of the system remains strictly Gaussian for that kind of multiplicative noise. Yet, since the multiplicative noise in the frictional damping alters the energy of the wave, the PDF becomes highly non-Gaussian.

The situation changes if the undamped trajectory does not coincide with the circular lines of constant strength of the multiplicative stochastic forcing. This situation occurs if non-identical additive noise amplitudes $\sigma_r^A \neq \sigma_i^A$ are used, or if a steady Rossby wave forcing [$\mathbf{F} \neq (0, 0)^T$] is imposed on the governing equation (B.1). Then, the stationary PDF becomes slightly non-Gaussian even in the case of a stochastic basic state. Nonetheless, the effect of a stochastic frictional damping on the non-Gaussianity of the systems PDF is much stronger than the corresponding effect of a stochastic basic state (see Fig. 6).

REFERENCES

- Baur, F., 1947: *Musterbeispiele Europäischer Großwetterlagen*. Dietrich, a Wiesbaden.
- Baur, F., F. Hess, and P. Nagel, 1944: *Kalender der Großwetterlagen Europas 1881-1939*. Forschungsinstitut für langfristige Witterungsvorhersage, Bad Homburg.
- Borges, M. D. and P. D. Sardeshmukh, 1995: Barotropic Rossby-wave dynamics of zonally varying upper-level flows during northern winter. *J. Atmos. Sci.*, **52**, 3779-3796.

- Cash, B. A. and S. Lee, 2001: Observed nonmodal growth of the Pacific North American teleconnection pattern. *J. Climate*, **14**, 1017-1028.
- Chang, J. S. and G. Cooper, 1970: A practical difference scheme for Fokker Planck equations. *J. Comput. Phys.*, **6**, 1-16.
- Charney, J. G. and J. G. DeVore, 1979: Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.*, **36**, 1205-1216.
- Cheng, X. and J. M. Wallace, 1993: Cluster analysis of Northern Hemisphere wintertime 500-hpa height field. *J. Atmos. Sci.*, **50**, 2674-2696.
- Corti, S., F. Molteni, and T. N. Palmer, 1999: Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature*, **29**, 799-802.
- DeSwart, H. E., 1988: Low-order spectral models of the atmospheric circulation: A survey. *Acta Appl. Math.*, **11**, 49-96.
- Egger, J., 1981: Stochastically driven large-scale circulations with multiple equilibria. *J. Atmos. Sci.*, **38**, 2606-2618.
- Ewald, B., C. Penland, and R. Temam, 2004: Accurate integrations of stochastic climate models. *Mon. Wea. Rev.*, **132**, 154-164.
- Farrell, B. F. and P. J. Ioannou, 1996: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.*, **53**, 2025-2040.
- Farrell, T. D. B. F., 1995: A stochastically excited linear system as a model for quasigeostrophic turbulence: Analytic results for one and two-layer fluids. *J. Atmos. Sci.*, **52**, 2531-2547.
- Friedrich, R., S. Siebert, J. Peinke, S. Lück, M. Siefert, M. Lindemann, J. Raethjen, G. Deusch, and G. Pfister, 2000: Extracting model equations from experimental data. *Phys. Lett. A*, **271**, 217-222.
- Gardiner, C. W., 1985: *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Science, Second Edition*. Springer-Verlag. 442 pp.
- Ghil, M. and S. Childress, 1987: *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory and Climate Dynamics*. Springer Verlag. 485 pp.
- Ghil, M. and A. W. Robertson, 2002: "Waves" vs. "particles" in the atmosphere's phase space: A pathway to long-range forecasting? *PNAS*, **99**, 2493-2500.
- Gradisek, J., S. Siebert, R. Friedrich, and I. Grabec, 2000: Analysis of time series from stochastic processes. *Phys. Rev. E*, **62**, 3146-3155.
- Hansen, A. R. and A. Sutera, 1986: On the probability density distribution of planetary scale atmospheric wave amplitude. *J. Atmos. Sci.*, **43**, 3250-3265.
- Hart, J., 1979: Barotropic quasigeostrophic flow over anisotropic mountains. *J. Atmos. Sci.*, **36**, 1736-1746.
- Hasselmann, K., 1976: Stochastic climate models. Part I. Theory. *Tellus*, **28**, 473-484.
- Horsthemke, W. and R. L ef ever, 1984: *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology*. Springer-Verlag. 318 pp.
- Hoskins, B. J., A. J. Simmons, and D. G. Andrews, 1977: Energy dispersion in a barotropic atmosphere. *Quart. J. Roy. Meteor. Soc.*, **103**, 553-567.
- Itoh, H. and M. Kimoto, 1996: Multiple attractors and chaotic itinerancy in a quasigeostrophic model with realistic topography: Implications for weather regimes and low-frequency variability. *J. Atmos. Sci.*, **53**, 2217-2231.
- Itoh, H. and M. Kimoto, 1997: Chaotic itinerancy with preferred transition routes appearing in an atmospheric model. *Phys. D*, **109**, 274-292.
- Itoh, H. and M. Kimoto, 1999: Weather regimes, low-frequency oscillations, and principal patterns of variability: A perspective of extratropical low frequency variability. *J. Atmos. Sci.*, **56**, 2684-2705.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.
- Khasminskii, R. Z., 1966: A limit theorem for the solutions of differential equations with random right hand side. *Theor. Prob. Appl.*, **11**, 390-406.
- Kimoto, M. and M. Ghil, 1993a: Multiple flow regimes in the northern hemisphere winter. Part I: Methodology and hemispheric regimes. *J. Atmos. Sci.*, **50**, 2625-2643.
- Kimoto, M. and M. Ghil, 1993b: Multiple flow regimes in the northern hemisphere winter. Part II: Sectorial regimes and preferred transitions. *J. Atmos. Sci.*, **50**, 2645-2672.
- Kloeden, P. and E. Platen, 1992: *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag. 632 pp.
- Landa, P. S. and P. V. E. McClintock, 2000: Changes in the dynamical behavior of nonlinear systems induced by noise. *Physics Reports*, **323**, 1-80.
- Legras, B. and M. Ghil, 1985: Persistent anomalies, blocking and variations in atmospheric predictability. *J. Atmos. Sci.*, **42**, 433-471.
- Lin, J. W. B. and J. D. Neelin, 2002: Considerations for stochastic convective parameterization. *J. Atmos. Sci.*, **59**, 959-975.
- Majda, A. J., I. Timofeyev, and E. Vanden Eijnden, 1999: stochastic climate prediction. *PNAS*, **96**, 14687-14691.
- Majda, A. J., I. Timofeyev, and E. Vanden Eijnden, 2003: Systematic strategies for stochastic mode reduction in climate. *J. Atmos. Sci.*, **60**, 1705-1722.
- Mo, K. and M. Ghil, 1988: Cluster analysis of multiple planetary flow Models for regimes. *J. Geophys. Res.*, **93**, 10927-10952.
- Molteni, F., S. Tibalti, and T. N. Palmer, 1990: Regimes in the wintertime circulation over northern extratropics. I. Observational evidence. *Quart. J. R. Met. Soc.*, **116**, 31-67.
- Monahan, A. H., J. C. Fyfe, and L. Pandolfo, 2003: The vertical structure of wintertime climate regimes of the northern hemisphere extratropical atmosphere. *J. Climate*, **16**, 2005-2021.
- Monahan, A. H., L. Pandolfo, and J. C. Fyfe, 2001: The preferred structure of variability of the Northern Hemisphere atmospheric circulation. *Geophys. Res. Lett.*, **28**, 1019-1022.
- Neelin, J. D. and W. Weng, 1999: Analytical prototypes for ocean-atmosphere interaction at midlatitudes. Part I. Coupled feedbacks as a sea surface temperature dependent stochastic process. *J. Climate*, **12**, 697-721.
- Newman, M., P. D. Sardeshmukh, and C. Penland, 1997: Stochastic forcing of wintertime extratropical flow. *J. Atmos. Sci.*, **54**, 435-455.
- Newman, M., P. D. Sardeshmukh, C. R. Winkler, and J. S. Whitaker, 2003: A study of subseasonal variability. *Mon. Wea. Rev.*, **31**, 1715-1732.
- Pandolfo, L., 1993: Observational aspects of the low-frequency intraseasonal variability of the atmosphere in middle latitudes. *Advances in Geophysics*, **34**, 93-174.
- Papanicolaou, G. and W. Kohler, 1974: Asymptotic theory of mixing stochastic differential equations. *Commun. Pure Appl. Math.*, **96**, 641-668.
- Park, B. T. and V. Petrosian, 1996: Fokker-Planck equations of stochastic acceleration: A study of numerical methods. *The Astrophysical Journal Supplement Series*, **103**, 255-267.
- Paul, W. and J. Baschnagel, 1999: *Stochastic Processes: From Physics to Finance*. Springer-Verlag. 231 pp.
- Penland, C., 1989: Random forcing and forecasting using principal oscillation pattern analysis. *Mon. Wea. Rev.*, **117**, 2165-2185.
- Penland, C., 1996: A stochastic model of IndoPacific sea surface temperature anomalies. *Phys. D*, **98**, 543-558.
- Penland, C. and L. Matrosova, 1994: A balance condition for stochastic numerical models with application to El Ni o - the Southern Oscillation. *J. Climate*, **7**, 1352-1372.
- Penland, C. and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999-2024.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in Fortran, The Art of Scientific Computing*. Cambridge University Press. 963 pp.
- Sardeshmukh, P. C. Penland, and M. Newman, 2001: Rossby waves in a fluctuating medium. In *Progress in Probability*, Vol. 49: *Stochastic Climate Models*, Imkeller, P. and von Storch, J.-S., editors. Birkh user Verlag, Basel.
- Sardeshmukh, P. D., G. P. Compo, and C. Penland, 2000: Changes of probability associated with El Ni o. *J. Climate*, **13**, 4268-4286.

- Sardeshmukh, P. D. and B. J. Hoskins, 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228-1251.
- Siegert, S., R. Friedrich, and J. Peinke, 1998: Analysis of data sets of stochastic systems. *Phys. Lett. A*, **243**, 275-280.
- Smyth, P., K. Ide, and M. Ghil, 1999: Multiple regimes in northern hemisphere height fields via mixture model clustering. *J. Atmos. Sci.*, **56**, 3704-3732.
- Sreenivasan, K. R., 1999: Fluid turbulence. *Rev. Mod. Phys.*, **71**, 383-395.
- Sreenivasan, K. R. and R. A. Antonia, 1997: The phenomenology of smallscale turbulence. *Annu. Rev. Fluid Mech.*, **29**, 435-472.
- Sura, P., 2002: Noise induced transitions in a barotropic β -plane channel. *J. Atmos. Sci.*, **59**, 97-110.
- Sura, P., 2003: Stochastic analysis of Southern and Pacific Ocean sea surface winds. *J. Atmos. Sci.*, **60**, 654-666.
- Sura, P. and J. J. Barsugli, 2002: A note on estimating drift and diffusion parameters from timeseries. *Phys. Lett. A*, **305**, 304-311.
- Sura, P. and S. T. Gille, 2003: Interpreting wind-driven Southern Ocean variability in a stochastic framework. *Journal of Marine Research*, **61**, 313-334.
- Sutera, A., 1986: Probability density distribution of large-scale atmospheric flow. *Adv. Geophys.* 227-249.
- van Bebber, W. J., 1891: Die Zugbahnen barometrischer Minima. *Meteorol. Z.*, **8**, 361-366.
- Weisheimer, A., D. Handorf, and K. Dethloff, 2001: On the structure and variability of atmospheric regimes in coupled climate models. *Atmospheric Science Letters*, **2**, 72-80.
- Wiin-Nielsen, A., 1979: Steady states and stability properties of a low-order barotropic system with forcing and dissipation. *Tellus*, **31**, 375-386.
- Winkler, C. R., M. Newman, and P. D. Sardeshmukh, 2001: A linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill. *J. Climate*, **14**, 4474-4493.
- Yoden, S., 1985a: Bifurcation properties of a quasi-geostrophic, barotropic, low-order model with topography. *J. Meteor. Soc. Japan*, **63**, 535-546.
- Yoden, S., 1985b: Multiple stable states of quasi-geostrophic barotropic flow over sinusoidal topography. *J. Meteor. Soc. Japan*, **63**, 1031-1044.